

Claims Description**Method and apparatus for pattern recognition employing the hidden Markov model**Patent Number: US5608841

Publication date: 1997-03-04

Inventor(s): TSUBOKA EIICHI (JP)

Applicant(s): MATSUSHITA ELECTRIC IND CO LTD (JP)

Requested Patent: JP5333898

Application Number: US19930071656 19930603

Priority Number(s): JP19920142399 19920603

IPC Classification: G10L5/06 ; G10L9/00

EC Classification: G10L15/14M ; G10L15/14M1

Equivalents: JP2795058B2

Abstract

A method and apparatus for estimating parameters of a new hidden Markov model (HMM) applicable to pattern recognition such as recognition of speech signals which are time series signals, and a method and apparatus for pattern recognition employing this HMM. It is easily applicable to speech and other time series signals. In particular, the pattern recognition degree of a time series observation vector signal y , received from an information source, is calculated by using function values $u(y,1)$, $u(y,2)$, ..., $u(y,M)$ and occurrence probabilities of signals C_1, C_2, \dots, C_M which are composed of set C , where $u(y,m)$ is the image into which the pair (C_m, Y) is mapped, $u(y,m) \in U$ ($U = [a,b]$, $a, b \in \text{epsilon } R$ and 0)

Data supplied from the esp@cenet database - I2

BEST AVAILABLE COPY

Specimen Received
319801975

E605t

(19)日本国特許庁 (JP)

(12) 公開特許公報(A)

(11)特許出願公開番号

特開平5-333898

(43)公開日 平成5年(1993)12月17日

(51) Int.Cl.⁶

識別記号 廈内整理番号
535 8842-5H
521 C 8842-5H

F L

技術表示箇所

(22)出願日 平成4年(1992)6月3日

(71)出願人 000005821

松下電器産業株式会社

大阪府門真市大字門真1006番地

(72)発明者 坪香 英

大阪府門真市大字門真1006番地 松下電器
産業株式会社内

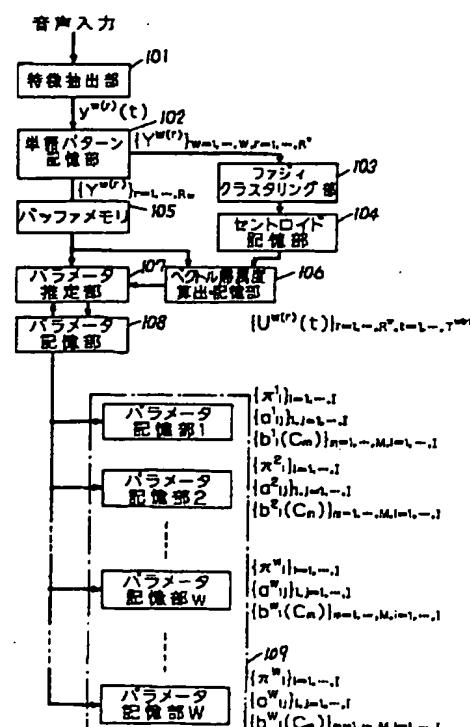
(74)代理人 弁理士 小鍛治 明 (外2名)

(54)【発明の名称】 時系列信号処理装置

(57) 【要約】

【目的】 HMMに基づく時系列パターンの処理装置であって、精度を落とさずに演算量を削減する。

【構成】 HMMの各状態における観測ベクトルの発生確率の対数値を、該状態における各クラスタの発生確率の対数値の荷重和とするものであって、該荷重和は前記入力ベクトルの前記クラスタに対する帰属度に関連した量を荷重係数とする。



【特許請求の範囲】

【請求項1】訓練ベクトル集合をクラスタリングして得られた各クラスタの発生確率を記憶するクラスタ発生確率記憶手段と、入力ベクトルの前記各クラスタに対する帰属度ベクトルを算出する帰属度ベクトル算出手段と、前記各クラスタ発生確率の対数値の、前記入力ベクトルに対する前記帰属度による、荷重和を算出する荷重和算出手段とを備え、該荷重和を前記入力ベクトルの発生確率または確率密度の対数値とすることを特徴とする時系列信号処理装置。

【請求項2】請求項1記載の時系列信号処理装置により得られるベクトルの発生確率に基づき、モデル化すべきベクトル系列からなるパターン集合の発生確率が最大になるように、該モデルのパラメータを推定するパラメータ推定手段を含むことを特徴とする隠れマルコフモデル作成装置。

【請求項3】請求項1記載の時系列信号処理装置により得られるベクトルの発生確率に基づき、モデル化すべきベクトル系列からなるパターン集合の発生確率が最大になるように、該モデルのパラメータを推定し、そのパラメータを記憶するモデルパラメータ記憶手段を備えたことを特徴とする隠れマルコフモデル記憶装置。

【請求項4】請求項3記載のモデルパラメータ記憶手段に記憶されている隠れマルコフモデルの、入力ベクトル系列に対する尤度を算出することを特徴とする尤度計算装置。

【請求項5】請求項3記載のモデルパラメータ記憶手段を認識単位毎に設け、入力ベクトル系列に対する前記モデル毎の尤度を計算する請求項4記載の尤度計算手段と、該尤度計算手段によって得られた前記モデル毎の尤度を比較し、その最大値を与えるモデルに対応した認識単位を認識結果とすることを特徴とする認識装置。

【発明の詳細な説明】

【0001】

【産業上の利用分野】音声認識等のパターン認識に適用可能な新しいHMM(Hidden Markov Model、隠れマルコフモデル)、そのパラメータの推定方法、該HMMを用いたパターンの認識方法、及び装置に関する。

【0002】

【従来の技術】本発明は一般的の時系列信号に適用可能なものであるが、説明の便宜のために、以下、音声認識を例に説明する。

【0003】先ずHMMを用いた音声認識装置について説明する。(図2)は、HMMを用いた音声認識装置のブロック図である。201は音声分析部であって、入力音声信号をフィルタバンク、フーリエ変換、LPC分析等の周知の方法により、一定時間間隔(フレームと呼ぶ)例えば10 msec毎に特徴ベクトルに変換する。従つて、入力音声信号は特徴ベクトルの系列 $Y = (y(1), y(2), \dots, y(T))$ に変換される。Tはフレーム数であ

る。202はコードブックと呼ばれるもので、ラベル付けされた代表ベクトルを保持している。203はベクトル量子化部であって、前記ベクトル系列Yのそれぞれのベクトルをそれに最も近い前記コードブックに登録されている代表ベクトルに対応するラベルに置き換えるものである。204はHMM作成部であって、訓練データから認識語乗たる各単語に対応するHMMを作成するものである。即ち、単語wに対応するHMMを作るには、まず、HMMの構造(状態数やそれら状態の間に許される遷移規則)を適当に定め、然る後に前記の如くして単語wを多数回発声して得られたラベル系列から、それらラベル系列の発生確率が出来るだけ高くなるように、前記モデルにおける状態遷移確率や状態の遷移に伴って発生するラベルの発生確率を求めるものである。205はHMM記憶部であって、このようにして得られたHMMを各単語毎に記憶するものである。206は尤度計算部であって、認識すべき未知入力音声のラベル系列に対し、前記HMM記憶部205に記憶されているそれぞれのモデルの前記ラベル系列に対する尤度を計算するものである。207は比較判定部であって尤度計算部206で得られた前記それぞれのモデルの尤度の最大値を与えるモデルに対応する単語を認識結果として判定するものである。

【0004】(図3)は離散確率分布型HMMの一例である。状態 q_i から状態 q_j への遷移確率 a_{ij} と、状態 q_i におけるラベル C_n の発生確率 $b_i(C_n)$ が定義されている。観測されたラベル系列の、このHMMから発生する確率は、Forward-Backward法やViterbi法により求められる。

【0005】HMMによる認識は具体的には次のようにして行われる。即ち、未知入力に対して得られたラベル系列を $O = (o(1), o(2), \dots, o(T))$ 、単語wに対応したモデルを λ^* とし、モデル λ^* により発生される長さTの任意の状態系列を $X = (x(1), x(2), \dots, x(T))$ とするとき、 λ^* のラベル系列Oに対する尤度は

【厳密解】

【0006】

【数1】

$$L_1(w) = \sum_x P(O, X | \lambda^*)$$

【0007】【近似解】

【0008】

【数2】

$$L_2(w) = \max_x [P(O, X | \lambda^*)]$$

【0009】または、対数をとって

【0010】

【数3】

$$L_3(w) = \max_x [\log P(O, X | \lambda^*)]$$

【0011】等で定義される。ここで、 $P(x, y | \lambda^w)$ は、モデル λ^w における x, y の同時確率である。

【0012】従って、例えば、(数1)を用いれば

【0013】

【数4】

$$w^* = \operatorname{argmax}_w [L_1(w)]$$

【0014】とするとき、 w^* が認識結果となる。(数2)、(数3)を用いるときも同様である。

$$P(O, X | \lambda) = \pi_{x(1)} \prod_{t=1}^T a_{x(t)} x(t+1) \prod_{t=1}^T b_{x(t)} (o(t))$$

【0017】と定義出来る。ここで $\pi_{x(1)}$ は状態 $x(1)$ の初期確率である。また、 $x(T+1) = 1+1$ は最終状態であって、如何なるラベルも発生しないとする。

【0018】この例では入力の特徴ベクトル $y(t)$ をラベル $o(t)$ に変換したが、各状態におけるラベルの発生確率の代りに特徴ベクトル $y(t)$ をそのまま用い、各状態において特徴ベクトル $y(t)$ の確率密度関数を与える方法もある。このときは(数5)における前記ラベル $o(t)$ の状態 q_i における発生確率 $b_i(o(t))$ の代わりに特徴ベクトル $y(t)$ の確率密度 $b_i(y(t))$ を用いることになる(以後、 z がラベルのときは $b_i(z)$ は z が状態 i において生じる確率、 z がベクトルのときは $b_i(z)$ は状態 i における z の確率密度を意味するものとする)。このときは、前記(数1)、(数2)、(数3)は次のようになる。

【厳密解】

【0019】

【数6】

$$L_1'(w) = \sum_x P(Y, X | \lambda^w)$$

【0020】[近似解]

【0021】

【数7】

$$L_2'(w) = \max_k [P(Y, X | \lambda^w)]$$

【0022】または、対数をとれば次式が得られる。

【0023】

【数8】

$$L_3'(w) = \max_x [\log P(Y, X | \lambda^w)]$$

【0024】以上、何れの方式を用いるにしても最終的な認識結果は、それぞれの単語 w に対してHMM λ^w を $w = 1 \sim W$ について準備しておけば、入力音声信号 Y に対して最大の尤度を与える λ^w に対応する w が認識結果となる。

【0025】

【発明が解決しようとする課題】前記従来例において、入力特徴ベクトルをラベルに変換するものを離散確率分布HMM、入力特徴ベクトルをそのまま用いるものを

【0015】 $P(O, X | \lambda)$ は次のようにして求められる。いま、HMM λ の状態 q_i ($i = 1 \sim I$)毎に、ラベル o の発生確率 $b_i(o)$ と状態 q_i ($i = 1 \sim I$)から状態 q_j ($j = 1 \sim I+1$)への遷移確率 a_{ij} が与えられているとき、状態系列 $X = (x(1), x(2), \dots, x(T+1))$ とラベル系列 $O = (o(1), o(2), \dots, o(T))$ のHMM λ から発生する同時確率は

【0016】

【数5】

$$P(O, X | \lambda) = \pi_{x(1)} \prod_{t=1}^T a_{x(t)} x(t+1) \prod_{t=1}^T b_{x(t)} (o(t))$$

統確率分布HMMと呼ぶことにする。このとき、これら両者の特徴は次のようにある。

【0026】離散確率分布HMMは、入力ラベル系列に対するモデルの尤度の計算において、各状態でのラベル C_m ($m = 1, \dots, M$)の発生確率 $b_i(C_m)$ はラベルに関連して予め記憶されている記憶装置から読み出すことで実行できるから計算量が非常に少ないと言う利点がある反面、量子化に伴う誤差のため、認識精度が悪くなると言う欠点がある。これを避けるためにラベル数(クラスタ数) M を多くする必要があるが、その増加に伴ってモデルを学習するために必要な学習パターン数が膨大になる。学習パターン数が不十分な場合は、前記 $b_i(C_m)$ が頻繁に0になることがあり、正しい推定が出来なくなる。例えば、次のようなことが生じる。

【0027】コードブックの作成は、認識すべき全ての単語について多数の話者の発聲音声を特徴ベクトル系列に変換し、この特徴ベクトルの集合をクラスタリングし、それぞれのクラスタにラベリングすることによって行われる。それぞれのクラスタは、セントロイドと呼ばれるそのクラスタの代表ベクトルを持ち、通常これは各々のクラスタに分類されたベクトルの期待値である。コードブックは、これらセントロイドを前記ラベルで検索可能な形で記憶したものである。

【0028】いま、前記認識語彙の中に、例えば「大阪」と言う単語があって、これに対応するモデルを作る場合を考える。多数話者が発声した単語「大阪」に対応する音声サンプルが特徴ベクトル列に変換され、各々の特徴ベクトルが前記セントロイドと比較され、最近隣のセントロイドがその特徴ベクトルの量子化されたものとなり、それに対応するラベルがその特徴ベクトルの符号化出力となる。このようにして、前記「大阪」に対する各々の音声サンプルは、ラベル系列に変換される。得られたラベル系列から、それらラベル系列に対する尤度が最大になるようにHMMのパラメータを推定することにより、単語「大阪」に対応するモデルが出来上がる。この推定には周知のBaum-Welch法等が用いられ得る。

【0029】この場合、前記コードブックにあるラベルの中には、単語「大阪」に対応する学習ラベル系列の中には含まれていないラベルが有り得る。このような場合

は、この含まれていないラベルの発生確率は「大阪」に対応するモデルにおいては学習の過程で“0”と推定されてしまう。従って、認識の時に発声される「大阪」と言う単語が変換されたラベル系列の中に、たまたま前記「大阪」のモデルの作成に用いたラベル系列には含まれていないラベルが存在することは十分有り得ることであり、その場合は、この認識時に発声された「大阪」のラベル系列が前記「大阪」のモデルから発生する確率は“0”になってしまう。ところが、このような場合でも、ラベルとしては異なっていても、ラベルに変換される前の特徴ベクトルの段階ではモデルの学習に用いた音声サンプルとかなり近く、ベクトルの段階で見れば十分「大阪」と認識されても良い場合がある。もともと同じ単語を発声しているのであるからベクトルのレベルでは似通っているにも関わらず、ラベルのレベルでは僅かの差で全く異なったラベルに変換されてしまうということは十分起り得るのであって、このようなことが認識精度に悪影響を及ぼすことは容易に想像がつく。クラスタ数Mが増加する程、訓練データ数が少ない程このような問題は頻繁に生じることになる。

【0030】この欠点を除去するためには、訓練集合には現れてこない（含まれていない）ラベルに対して、平滑化や補完を行う等の工夫が必要となる。「結び」と呼ばれる概念を用いてパラメータ数を減少させる工夫をはじめとして、0確率が推定される場合はそれを0にせず微小量に置き換えたり、ファジイベクトル量子化等のようにクラスタの境界をぼかしたりする方法等、平滑化や補完を行う方法が種々提案されている。中でもファジイベクトル量子化に基づくHMMは、ヒューリスティックな要素が少なく、理論的にもすっきりしていて、アルゴリズミックに実現できる方法であるが、従来提案されているものは数学的な意味で近似的なものであった。

【0031】

【課題を解決するための手段】訓練ベクトル集合をクラスタリングして得られた各クラスタの発生確率を記憶するクラスタ発生確率記憶手段と、入力ベクトルの前記各クラスタに対する帰属度ベクトルを算出する帰属度ベクトル算出手段と、前記各クラスタ発生確率の対数値の、前記入力ベクトルに対する前記帰属度による、荷重和を算出する荷重和算出手段とを備え、該荷重和を前記入力ベクトルの発生確率の対数値とすることを特徴とする構成である。

【0032】

【作用】クラスタ発生確率記憶手段により訓練ベクトル集合をクラスタリングして得られた各クラスタの発生確率を記憶し、帰属度ベクトル算出手段により入力ベクトルの前記各クラスタに対する帰属度ベクトルを算出し、荷重和算出手段により前記各クラスタ発生確率の対数値の、前記入力ベクトルに対する前記帰属度による荷重和を算出し、該荷重和を前記入力ベクトルの発生確率の対

数値とするものである。

【0033】

【実施例】ここで、以後用いるべき記号の定義をまとめておく。簡単のために、誤解を生じない限り、状態 q_i, q_j 等は単に i, j 等と表記することにする。また、モデルの学習は1つの単語について行う場合を述べることとし、モデル間で区別する必要のある場合は、モデルに対応する番号をパラメータの右肩に添字として付加することとし、通常はこれを省くものとする。

【0034】 $i = 1, 2, \dots, I + 1$: 第*i*番の状態

$[a_{ij}]$: 遷移マトリクス

a_{ij} : 状態*i*から状態*j*への遷移確率

r : 作成すべきモデルに対する訓練パターン番号 ($r = 1, \dots, R$)

$y^{(r)}(t)$: 訓練パターン*r*の第*t*フレームにおける観測ベクトル

$o^{(r)}(t)$: 訓練パターン*r*の第*t*フレームにおける観測ラベル

$b_i(y^{(r)}(t))$: 訓練パターン*r*のフレーム*t*の観測ベクトル $y^{(r)}(t)$ の状態*i*における確率密度

$Y^{(r)} = (y^{(r)}(1), y^{(r)}(2), \dots, y^{(r)}(T^{(r)}))$: 訓練パターン*r*のベクトル系列 (ただし、 $r = 1, 2, \dots, R$)

$O^{(r)} = (o^{(r)}(1), o^{(r)}(2), \dots, o^{(r)}(T^{(r)}))$: 訓練パターン*r*のラベル系列 (ただし、 $r = 1, 2, \dots, R$)

$X^{(r)} = (x^{(r)}(1), x^{(r)}(2), \dots, x^{(r)}(T^{(r)}), x^{(r)}(T^{(r)} + 1))$: $X^{(r)}$ または $O^{(r)}$ に対応する状態系列

$x^{(r)}(t)$: 単語*w*に対する第*r*番の訓練パターンの第*t*フレームにおける状態

$T^{(r)}$: 単語*w*に対する第*r*番の訓練パターンのフレーム数

$\lambda_i = [(a_{ij})_{j=1, \dots, I+1}, (b_i(C_m))_{m=1, \dots, M}]$: 状態*i*のパラメータの集合

S_m : 第*m*クラスターに含まれるベクトル集合

C_m : 第*m*番のクラスター名 (ラベル)

c_m : C_m 従って S_m のセントロイド (S_m の重心ベクトル)

$b_i(C_m)$: 訓練パターン*r*のフレーム*t*の観測ラベル C_m の状態*i*における発生確率

$\lambda = \{\lambda_i\}$: 全パラメータの集合 (λ をパラメータとするモデルをモデル λ とも呼ぶ)

$P(Y|\lambda)$: 観測ベクトル系列 Y がモデル λ から発生する確率密度

$P(O|\lambda)$: 観測ラベル系列 O がモデル λ から発生する確率

π_i : 状態*i*が $t = 1$ で生じる確率

先ず、離散確率分布HMMを学習する方法について述べる。

【0035】離散確率分布HMMでは、ベクトル系列からなる訓練パターンを構成する各ベクトルはベクトル量

子化 (Vector Quantize) され、ラベル系列に変換されるそこで、先ず、VQについて簡単に説明する。

【0036】 (1) コードブックの作成

取り扱うべき信号の多数の訓練集合から特徴抽出を行い、特徴ベクトル集合を得（前記訓練パターンを構成するベクトル等）、この特徴ベクトル集合をクラスタリングし、M個のクラスタ S_1, \dots, S_m と各々のクラスタの重心ベクトル（セントロイド） c_1, \dots, c_m を得、 $m = 1, \dots, M$ について、 m により参照可能な形で c_m を記憶する。これをコードブックという。

【0037】 (2) ベクトル量子化・符号化

符号化すべきベクトル y をベクトル c_1, \dots, c_m の何れかで近似することをベクトル量子化と言い、 y が c_m に量子化されたとき、 y をラベル C_m に置き換えることを y を C_m に符号化すると言う。ベクトル系列 $y(1), \dots, y(T)$ は通常次のような方法でラベル系列 $o(1), \dots, o(T)$ に変換される。

【0038】 符号化すべき信号から (1) で行ったのと同様の方法により特徴抽出を行い、特徴ベクトル列 $y(1), y(2), \dots, y(T)$ を得たとする。 $y(t)$ と c_m の距離を $d(y(t), c_m)$ とするとき、

【0039】

【数9】

$$o(t) = \underset{\{1, 2, \dots, m\}}{\operatorname{argmin}} [d(y(t), c_m)]$$

【0040】 $o(t)$ を $y(t)$ の符号化出力であるとする ($o(t) \in \{1, 2, \dots, M\}$)。 $d(y(t), c_m)$ としては、ユークリッドノルム等が用いられる。

【0041】 前記クラスタリングには、フルサーチクラスタリング、バイナリーツリークラスタリング等の方法があり、それぞれについて種々の方法が考えられている。例えば、フルサーチクラスタリングの一つとして、次のようにして行うものがある。訓練ベクトル集合を v_1, v_2, \dots, v_N とする。

【0042】 (1) 任意にM個のベクトル c_1, \dots, c_m を定める。

(2) $m = 1, \dots, M$ について

【0043】

【数10】

$$S_m = \{v_n \mid d(v_n, c_m) < d(v_n, c_{m'}), \forall m' \neq m\}$$

【0044】 を計算する。

(3) $m = 1, \dots, M$ について、クラスタ S_m のセントロイドを求め、各々を各クラスタの新たなセントロイドとして c_m を更新する。即ち、

【0045】

【数11】

$$c_m = \frac{1}{|S_m|} \sum_{n \in S_m} v_n$$

【0046】 ただし、本式において、 $|S_m|$ は S_m の要素数を意味するものとする。

(4) 収束条件を検査し、それが満たされていれば完了し、満たされていない場合はステップ (2) に戻る。

【0047】 収束条件としては、a) 歪の減少率がそれに関する設けた閾値 ϵ 以下になった場合、b) ステップ (2) ~ (4) の繰り返しの回数がそれに関する設けた制限 I に到達した場合等が採用される。a) については例えば次のように実行され得る。即ち、ステップ (2) ~ ステップ (4) の第 i 番の繰り返しにおいて得られる歪量

【0048】

【数12】

$$D(i) = \sum_{n=1}^N \sum_{v_n \in S_m} d(c_m, v_n)$$

【0049】 を定義し、予め定めた小さな数 ϵ に対して $\epsilon > |D(i-1) - D(i)| / D(i)$ となれば、収束したと見なす等である。

【0050】 以上説明したクラスタリングは、後に説明するファジィクラスタリングがソフトクラスタリングと呼ばれるのに対してハードクラスタリングと呼ばれることがある。

【0051】 離散HMM作成の問題は、ある認識単位（単語等）に対し、準備された $r = 1 \sim R$ の訓練パターンから尤度関数 $P(O^{(1)}, O^{(2)}, \dots, O^{(R)} | \lambda)$ を最大にする前記HMMを規定するパラメータ λ を推定することである。

【0052】 $O^{(r)}$ が互いに独立であるとすれば、前記尤度関数は

【0053】

【数13】

$$P(O^{(1)}, O^{(2)}, \dots, O^{(R)} | \lambda)$$

$$= \prod_{r=1}^R P(O^{(r)} | \lambda)$$

$$= \prod_{r=1}^R \left(\sum_{X^{(r)}} P(O^{(r)}, X^{(r)} | \lambda) \right)$$

$$= \sum_{X^{(1)}} \dots \sum_{X^{(R)}} \prod_{k=1}^R P(O^{(k)}, X^{(k)} | \lambda)$$

【0054】 で与えられる。ここで、次の補助関数 $Q(\lambda, \lambda')$ を定義する。

【0055】

【数14】

$$Q(\lambda, \lambda')$$

$$= \sum_{X^{(1)}} \dots \sum_{X^{(R)}} \prod_{k=1}^R P(O^{(k)}, X^{(k)} | \lambda) \prod_{k=1}^R V_{k-1} \quad \textcircled{②}$$

$$\log \frac{\prod_{k=1}^R P(O^{(k)}, X^{(k)} | \lambda)}{\prod_{k=1}^R V_{k-1}} \quad \textcircled{③}$$

【0056】 このとき、次のことが言える。「 $Q(\lambda, \lambda') \geq Q(\lambda, \lambda)$ なら、 $P(O^{(1)}, \dots, O^{(R)} | \lambda') \geq P(O^{(1)}, \dots, O^{(R)} | \lambda)$ 」

$(O^{(1)}, \dots, O^{(R)} | \lambda)$ であつて、等号は $\lambda' = \lambda$ の時に成り立つ。」故に、

【0057】

【数15】

$$\lambda' = \underset{\lambda}{\operatorname{argmax}} [Q(\lambda, \lambda')]$$

【0058】を求めることが出来れば、 $\lambda' \rightarrow \lambda$ として (数15) を繰り返し適用することによって、 λ は $P(O^{(1)}, \dots, O^{(R)} | \lambda)$ の停留点、即ち、 $P(O^{(1)}, \dots, O^{(R)} | \lambda)$

$$= P(O^{(1)}, \dots, O^{(R)} | \lambda) \sum_{r=1}^R \frac{1}{P(O^{(r)}, X^{(r)} | \lambda)} \sum_{x^{(r)}} P(O^{(r)}, X^{(r)} | \lambda) \log P(O^{(r)}, X^{(r)} | \lambda)$$

【0061】前述の説明から、 $Q(\lambda, \lambda')$ を λ' の関数と見なして $Q(\lambda, \lambda') > Q(\lambda, \lambda)$ なる λ' を見出せば、それは λ の更新されたものとなるが、 $P(O^{(1)}, \dots, O^{(R)} | \lambda)$ は λ' に関しては一定値となるから、これを取り

$$Q'(\lambda, \lambda') = Q(\lambda, \lambda') / P(O^{(1)}, \dots, O^{(R)} | \lambda) \\ = \sum_{r=1}^R h^{(r)} \sum_{x^{(r)}} P(O^{(r)}, X^{(r)} | \lambda) \log P(O^{(r)}, X^{(r)} | \lambda')$$

【0063】とするとき、前記求むべき λ' を見出すことは $Q'(\lambda, \lambda') > Q'(\lambda, \lambda)$ なる λ' を見出すことと同様である。

$$\xi^{(r)}_{ij}(t) = P(O^{(r)}, X^{(r)}(t-1) = i, X^{(r)}(t) = j | \lambda) \\ \tau^{(r)}_i(t) = P(O^{(r)}, X^{(r)}(t) = i | \lambda) = \sum_{j=1}^{I+1} \xi^{(r)}_{ij}(t) \\ h^{(r)} = 1 / P(O^{(r)} | \lambda) = 1 / \sum_{i=1}^I \tau^{(r)}_i(t)$$

とおけば、(数17) はさらに次のようになる。

【0066】

$$Q'(\lambda, \lambda') \\ = \sum_{r=1}^R h^{(r)} \sum_{x^{(r)}} P(O^{(r)}, X^{(r)} | \lambda) \\ \times \{ \log \pi_{x^{(r)}(t)}' + \sum_i \log a_{x^{(r)}(t) x^{(r)}(t+1)}' \\ + \sum_i \log b_{x^{(r)}}(o^{(r)}(t))' \} \\ = \sum_{r=1}^R h^{(r)} \sum_i \tau^{(r)}_i(t) \log \pi_i' \\ + \sum_{r=1}^R h^{(r)} \sum_i \sum_j \sum_{x^{(r)}} \xi^{(r)}_{ij}(t) \log a_{ij}' \\ + \sum_{r=1}^R h^{(r)} \sum_i \sum_j \tau^{(r)}_i(t) \sum_{n=1}^N \delta(o^{(r)}(t), C_n) \log b_i(C_n)'.$$

【0067】ここで、 $\delta(i, j)$ はいわゆるクロネッカーデルタで、 $i = j$ のとき $\delta(i, j) = 1$ 、 $i \neq j$ のとき $\delta(i, j) = 0$ である。各パラメータの再推定値は Lagrange の未定乗数法によって導かれる。

【0068】右辺第1項から π_i' について

【0069】

$(R | \lambda)$ の極大値または鞍点を与える点に収束することになる。 $P(O^{(1)}, \dots, O^{(R)} | \lambda)$ の変化率が予め定めた閾値以下になるまでこの操作を繰り返すことにより局所最適解が得られる。

【0059】この $Q(\lambda, \lambda')$ を用いてパラメータを推定する方法について説明する。(数14) を変形すれば、次式が得られる。

【0060】

【数16】

除いて

【0062】

【数17】

【0064】ここで

【0065】

【数18】

$\xi^{(r)}_{ij}(t) = P(O^{(r)}, X^{(r)}(t-1) = i, X^{(r)}(t) = j | \lambda)$

$$\tau^{(r)}_i(t) = P(O^{(r)}, X^{(r)}(t) = i | \lambda) = \sum_{j=1}^{I+1} \xi^{(r)}_{ij}(t)$$

$$h^{(r)} = 1 / P(O^{(r)} | \lambda) = 1 / \sum_{i=1}^I \tau^{(r)}_i(t)$$

【数19】

【数20】

$$\sum_{i=1}^I \pi_i' = 1$$

【0070】の条件のもとで最大化すれば π_i' の再推定値 π_i^* は

【0071】

【数21】

$$\pi_i = \sum_{r=1}^R h^{(r)} \tau_i(1).$$

【0072】右辺第2項から a_{ij}' について

【0073】

【数22】

$$\sum_{j=1}^J a_{ij}' = 1$$

【0074】なる条件のもとで最大化すれば a_{ij} の再推定値 a_{ij}^* は

【0075】

【数23】

$$a_{ij}^* = \frac{\sum_{r=1}^R h^{(r)} \sum_{t=1}^{T(r)} \xi^{(r)}_{ij}(t)}{\sum_{r=1}^R h^{(r)} \sum_{t=1}^{T(r)} \sum_{j=1}^J \xi^{(r)}_{ij}(t)}$$

$$b_i(C_n) = \frac{\sum_{r=1}^R h^{(r)} \sum_{t=1}^{T(r)} \tau^{(r)}_i(t) \delta(o^{(r)}(t), m)}{\sum_{n=1}^N \sum_{r=1}^R h^{(r)} \sum_{t=1}^{T(r)} \tau^{(r)}_i(t) \delta(o^{(r)}(t), m)}$$

$$= \frac{\sum_{r=1}^R h^{(r)} \sum_{t=y}^{T(r)} \sum_{i=1}^{T(r)} \tau^{(r)}_i(t) s_m}{\sum_{r=1}^R h^{(r)} \sum_{t=1}^{T(r)} \tau^{(r)}_i(t)}$$

【0080】ここで、 $\xi^{(r)}_{ij}(t)$, $\gamma^{(r)}_i(t)$ は次の
ように計算される。即ち、

$$\alpha^{(r)}_i(t) = P(o^{(r)}(1), \dots, o^{(r)}(t), x^{(r)}(t) = i | \lambda) \quad \oplus$$

$$\beta^{(r)}_i(t) = P(o^{(r)}(t+1), \dots, o^{(r)}(T) | x^{(r)}(t) = i, \lambda) \quad \oplus$$

【0082】とおけば、

【0083】

$$\xi^{(r)}_{ij}(t) = P(o^{(r)}, x^{(r)}(t-1) = i, x^{(r)}(t) = j | \lambda)$$

$$= \alpha^{(r)}_i(t-1) a_{ij} b_j(o^{(r)}(t)) \beta^{(r)}_j(t)$$

$$\tau^{(r)}_i(t) = \alpha^{(r)}_i(t) \beta^{(r)}_i(t)$$

【0084】であって、

【0085】

【数28】

$$\alpha^{(r)}_i(t) = \sum_j \alpha^{(r)}_i(t-1) a_{ij} b_j(o^{(r)}(t))$$

【0086】

【数29】

$$\beta^{(r)}_i(t) = \sum_j b_j(o^{(r)}(t)) a_{ij} \beta^{(r)}_j(t)$$

【0087】なる漸化式が成り立つから、 $\alpha^{(r)}_i(0) = 1$ としてパラメータ λ に適当な初期値を与える、 $t = 1 \sim T^{(r)} + 1$, $j = 1 \sim I + 1$ について (数28) に従つて $\alpha^{(r)}_j(t)$ を、 $\beta^{(r)}_{I+1}(T^{(r)} + 1) = 1$ として $t = T^{(r)} + 1 \sim 1$, $i = I \sim 1$ について (数29) に従つ【0076】右辺第3項から $b_i(C_n)$ について

【0077】

【数24】

$$\sum_{n=1}^N b_i(C_n) = 1$$

【0078】の条件のもとに最大化すれば、 $b_i(C_n)$ の再推定値 $b_i(C_n)^*$ は

【0079】

【数25】

【0081】

【数26】

【数27】

て $\beta^{(r)}_i(t)$ をそれぞれ順次計算して行けば、 (数27) が計算できる。

【0088】パラメータ推定の実際の計算手順は次のようになる。

(1) $L_1 = \infty$ (2) $i = 1 \sim I$, $j = 1 \sim I + 1$, $m = 1 \sim M$ について

$$\lambda_i = [\{\pi_{ij}\}_{j=1, \dots, I}, \{a_{ij}\}_{j=1, \dots, I+1}, \{b_i(C_n)\}_{n=1, \dots, N}]$$

に適当な初期値を与える。

【0089】(3) $r = 1 \sim R$, $t = 2 \sim T^{(r)}$, $i = 1 \sim I$, $j = 1 \sim I + 1$ について(数28) (数29) に従つて $\alpha^{(r)}_i(t)$, $\beta^{(r)}_i(t)$ を (数27) に従つて $\xi^{(r)}_{ij}(t)$, $\gamma^{(r)}_i(t)$ を求め

る。

【0090】(4) $i = 1 \sim I$ について

(数21) に従って π_i の再推定値 π_i^* を求める。

【0091】(5) $i = 1 \sim I, j = 1 \sim I + 1$ について

(数23) に従って a_{ij} の再推定値 a_{ij}^* を求める。

【0092】(6) $i = 1 \sim I, m = 1 \sim M$ について
(数25) に従って $b_i(C_m)$ の再推定値 $b_i(C_m)^*$ を求める。

【0093】(7) $i = 1 \sim I, j = 1 \sim I + 1, m = 1 \sim M$ について

$\pi_i = \pi_i^*$, $a_{ij} = a_{ij}^*$, $b_i(C_m) = b_i(C_m)^*$ なる代入を行うことによって、再推定されたパラメータ集合 $\lambda = \{\lambda_i\}$ を得る。

【0094】(8) step (7) で得たパラメータ集合 λ に対して

【0095】

【数30】

$$L_2 = \sum_{r=1}^R P(O^{(r)} | \lambda) = \sum_{r=1}^R \alpha^{(r)} \lambda_{r+1} (T^{(r)} + 1)$$

【0096】を計算する。

(9) $|L_1 - L_2| / L_1 > \epsilon$ ならば、 $L_2 = L_1$ とおいてステップ (4) へ、そうでなければ終了。

【0097】前記ステップ (9) における ϵ は収束の幅を決める適当に小さな正の数であって、その値は状況によって実用的な値が選ばれる。

【0098】以上のようにして、離散確率分布HMMが得られるが、前述のごとき欠点を有している。次に、前記離散確率分布HMMを基にしてファジイクラスタリング(ファジイベクトル量子化)によるHMM (FVQHMM) を説明する。

【0099】先ず、ファジイクラスタリングについて説明する。表記の簡単のために、学習に用いる全ての単語

$$u_{mn} = \frac{1}{N} \sum_{k=1}^N \frac{d_{mn}}{d_{kn}}$$

【0107】ここで、Fはいわゆるファジイネスを表し、 $1 < F$ である。 $F \rightarrow \infty$ のときは、 $m = 1, \dots, M$ について、 $u_{mn} \rightarrow 1/M$ になり、 $F \rightarrow 1$ のときは、

$$(d_{mn} / d_{kn})^{1/(F-1)} \rightarrow \begin{cases} 1 & m = n \\ 0 & m \neq n \end{cases}$$

$m = n$ のとき

$m \neq n, d_{mn} < d_{kn}$ のとき

$m \neq n, d_{mn} > d_{kn}$ のとき

【0109】であるから、

【0110】

【数36】

音声を構成するベクトルについて通し番号をつけ、 y_1, \dots, y_N とし、 y_n のクラスタ S_n への帰属度(メンバシップ値)を u_{mn} とする。通常のクラスタリング法では、あるベクトル y_n がクラスタ S_n に属している ($u_{mn} = 1$) か、属していないか ($u_{mn} = 0$) だけを認めるのに対して、ファジイクラスタリング法では、 y_n がいくつかのクラスタに異なる度合で帰属することを認めると言うものである。

【0100】具体的な方法の1つに、各クラスタ S_n のセントロイド(中心ベクトル、平均ベクトル)を c_n ($m = 1, \dots, M$) とし、 y_n とセントロイド c_n の非類似度を $d_{mn} = d(y_n, c_n)$ とするとき、

【0101】

【数31】

$$J = \sum_{m=1}^M \sum_{n=1}^N (u_{mn})^F d(y_n, c_m) \quad F > 1$$

【0102】を最小にする u_{mn} と c_n を見出すものがある。目的関数 J を c_n と u_{mn} に関して偏微分し、条件

【0103】

【数32】

$$\sum_{n=1}^N u_{mn} = 1$$

【0104】を用いれば、 J を局所的に最小化するための必要条件が次の(数30) (数31) のように示される。

【0105】

【数33】

$$c_n = \frac{\sum_{m=1}^M (u_{mn})^F y_n}{\sum_{n=1}^N (u_{mn})^F} \quad 1 \leq m \leq M.$$

【0106】

【数34】

$$1 \leq m \leq M, \quad 1 \leq n \leq N.$$

【0108】

【数35】

$m = n$ のとき

$m \neq n, d_{mn} < d_{kn}$ のとき

$m \neq n, d_{mn} > d_{kn}$ のとき

$$u_{mn} \rightarrow \begin{cases} 1 & d_{mn} \text{ が最小のとき} \\ 0 & d_{mn} \text{ が最小でないとき} \end{cases}$$

【0111】となる。即ち、Fが増加するにつれて y_n が何れのクラスタに属するかと言うことの曖昧性が増大

し、Fが1に近づくにつれて、 y_n の属するクラスタを一意に決定するいわゆるハードクラスタリングに近づく。

【0112】ファジィクラスタリングの実際の手順は次のようになる。

(1) 訓練ベクトル集合を適当に S_1, \dots, S_n に分割し、初期クラスタとする。

【0113】 u_{mn} を適当に初期化する。

(2) 各クラスタの平均ベクトル c_m を(数30)に従って求める。

【0114】(3) $y_n \neq c_m$ のとき、(数31)によつてステップ(2)の結果を用いて u_{mn} を更新する。 $y_n = c_m$ のときは、 $m=n$ のとき $u_{mn} = 1$ 、 $m \neq n$ のとき $u_{mn} = 0$ とおく。

【0115】(4) 収束条件を満足すれば処理を終了。

そうでなければステップ(2)へ戻る。

【0116】ステップ(4)における収束条件としては、上記繰り返し計算において、繰り返し回数をkとして更新される前のJを $J(k)$ 、更新された後のJを $J(k+1)$ とするとき、 $|J(k) - J(k+1)| / J(k+1)$ が予め定めた収束判定値 ϵ 以下になった場合、あるいは、適当に定めた繰り返し回数の上限をKとして $k = K$ となったときを収束、何れにも達しない場合を非収束とする等が考えられる。

【0117】以上の結果を用いてFVQHMMを構成する従来法は次のようなものであった。即ち、状態iにおける $y^{(r)}(t)$ の生起確率を $\omega_i(y^{(r)}(t))$ とすると

【0118】

【数37】

$$\omega_i(y^{(r)}(t)) = \frac{1}{\sum_{m=1}^M b_i(C_m) \sum_{k=1}^K u(y^{(r)}(t), C_k) \oplus \frac{1}{\sum_{m=1}^M b_i(C_m) \sum_{k=1}^K u(y^{(r)}(t), C_k) \oplus \frac{1}{\sum_{m=1}^M b_i(C_m) \sum_{k=1}^K u(y^{(r)}(t), C_k) \oplus \dots}}}$$

【0119】とおき、(数27)～(数29)における $b_i(o^{(r)}(t))$ を $\omega_i(y^{(r)}(t))$ で置き換え、(数25)を

$$b_i(C_m) = \frac{\sum_{r=1}^R \sum_{t=1}^{T(r)} u(y^{(r)}(t), C_m) \tau^{(r)}(t)}{\sum_{r=1}^R \sum_{t=1}^{T(r)} \tau^{(r)}(t)}$$

【0121】に置き換えたものである。この方法の問題点は、結果的に

【0122】

【数39】

$$\log \sum_m u(y^{(r)}(t), C_m) b_i(C_m) = \sum_m u(y^{(r)}(t), C_m) \log b_i(C_m)$$

【0123】という近似を行っていることである。即ち、後に述べるように(数37)の定義のもとに(数38)を導くには、(数39)でなければならぬ。ところが一般に凸関数の性質から

【0124】

【数40】

$$\log \sum_m u(y^{(r)}(t), C_m) b_i(C_m) \leq \sum_m u(y^{(r)}(t), C_m) \log b_i(C_m)$$

$$\omega_i(y^{(r)}(t)) = \prod_{m=1}^M \frac{b_i(C_m)}{\sum_{k=1}^K b_i(C_k)} \oplus \frac{1}{\sum_{m=1}^M b_i(C_m) \sum_{k=1}^K u(y^{(r)}(t), C_k) \oplus \dots}$$

即ち

$$\log \omega_i(y^{(r)}(t)) = \sum_{m=1}^M u(y^{(r)}(t), C_m) \log b_i(C_m)$$

【0125】であつて、等号が成立するのは、(1)はあるmについて $u(y^{(r)}(t), C_m) = 1$ 、その他のmについては $u(y^{(r)}(t), C_m) = 0$ の場合、(2) $b_i(C_m)$ がmに関わらずすべて等しい場合の何れかである。従つて、上の近似が良くなるのは前記Fが1に近い、即ち、クラスタリングが前記ハードクラスタリングに近い場合か、クラスタ数が少なくて、 $b_i(C_m)$ の値がmに関してあまり変動が無い場合に限られる。

【0126】本発明は以上の欠点を廃して、数学的に矛盾の無いFVQHMMを提供するものである。即ち前記 $\omega_i(y^{(r)}(t))$ を次式で定義するものである。

【0127】

【数41】

【0128】このようにおけば、(数19)の右辺第3項(Q_3 とおく)は、

【0129】

【数42】

$$\sum_{n=1}^N \delta(o^{(r)}(t), C_n) \log b_i(C_n)$$

$$Q_3 = \sum_{r=1}^R h^{(r)} \sum_i \sum_{n=1}^N r^{(r)}(t) \sum_{n=1}^N u(o^{(r)}(t), C_n) \log b_i(C_n)'.$$

【0132】となる。即ち、(数41)のように定義することにより、ハードクラスタリングにおける $\delta(o^{(r)}(t), C_n)$ を $u(o^{(r)}(t), C_n)$ に置き代えたことになる。このとき、前記説明から

【0133】

【数44】

$$u(o^{(r)}(t), C_n) \rightarrow \delta(o^{(r)}(t), C_n) \quad (F \rightarrow 1)$$

【0134】が言えるから、本発明によるファジィクラスタリングは、ハードクラスタリングの自然な拡張となっていることがわかる。

【0135】 $b_i(C_n)$ の再推定値は Q_3 を $b_i(C_n)'$ に關

$$0 = \frac{\partial}{\partial b_i(C_n)'} \frac{\sum_{r=1}^R Q_3 - \sum_{k=1}^N b_i(C_k)'}{\sum_{k=1}^N b_i(C_k)'} - \frac{1}{\sum_{k=1}^N b_i(C_k)'} \quad (4)$$

$$= \sum_{r=1}^R h^{(r)} \sum_i r^{(r)}(t) u(o^{(r)}(t), C_n) \frac{1}{b_i(C_n)'} - \theta$$

【0140】である。両辺を $b_i(C_n)'$ 倍し、 m について総和をとる。 $u(o^{(r)}(t), C_n)$ の m に関する総和は1であるということに注意すれば、

【0141】

【数47】

$$\theta = \sum_{n=1}^N \sum_{r=1}^R h^{(r)} \sum_{i=1}^{T(r)} r^{(r)}(t) u(o^{(r)}(t), C_n)$$

$$= \sum_{r=1}^R h^{(r)} \sum_{i=1}^{T(r)} r^{(r)}(t)$$

$$b_i(C_n)' = \frac{\sum_{r=1}^R h^{(r)} \sum_{i=1}^{T(r)} r^{(r)}(t) u(o^{(r)}(t), C_n)}{\sum_{r=1}^R h^{(r)} \sum_{i=1}^{T(r)} r^{(r)}(t)}$$

【0144】が得られる。これは、(数38)と同じ形をしているものであるが、(数41)のように $\omega_i(y^{(r)}(t))$ を定義することによって導かれるべきものであり、(数37)の定義からは導くことは出来ない。この意味で、 $\omega_i(y(t))$ を(数37)で定義し、 $b_i(C_n)$ の再推定値を(数38)で与えている従来例は(数39)の近似を行っていることになる。

【0145】本実施例では、観測ベクトル $y(t)$ から算出されるクラスタ C_n に対する荷重係数(帰属度) $u(y^{(r)}(t), C_n)$

【0130】の代わりに $\log \omega_i(y^{(r)}(t))$ を代入したものとなり

【0131】

【数43】

して

【0136】

【数45】

$$\sum_{n=1}^N b_i(C_n)' = 1$$

【0137】の条件のもとで最大化することにより得られる。この手順を少し詳しく述べれば次のようになる。

【0138】Lagrangeの未定乗数を θ とすれば、

【0139】

【数46】

【0142】となる。(数47)を(数46)に代入して整理すれば、

【0143】

【数48】

$(r)(t), C_n)$ はファジィクラスタリングにおけるメンバシップ値として説明したが、例えば、クラスタ C_n における $y(t)$ の条件付確率(密度)や前記メンバシップ値の n 乗(n は整数)等場合に応じて種々の定義が考えられる。この場合は、一般に、 $u(y^{(r)}(t), C_n)$ の m に関する総和は1ではなく、この条件を考慮しなければ、(数48)は(数49)のようになるのは明かである。

【0146】

【数49】

$$b_i(C_n) = \frac{\sum_{r=1}^R h^{(r)} \sum_{t=1}^{T(r)} T^{(r)}_i(t) u(o^{(r)}(t), C_n)}{\sum_{r=1}^R h^{(r)} \sum_{t=1}^{T(r)} T^{(r)}_i(t) \sum_{a=1}^n u(o^{(r)}(t), C_n)}$$

【0147】他のパラメータの推定式は、ハードクラスタリングの場合と同様であって、得られる結果も表記的には(数21) (数23)と全く同じである。但し、 α , β の計算式(数28) (数29)が(数50)のように変更され、(数27)における δ の計算式は(数51)のように変更される。

【0148】

$$\begin{aligned}\xi^{(r)}_{i,j}(t) &= \alpha^{(r)}_i(t-1) a_{ij} \omega_j(y^{(r)}(t)) \beta^{(r)}_j(t), \\ \tau^{(r)}_i(t) &= \alpha^{(r)}_i(t) \beta^{(r)}_i(t).\end{aligned}$$

【0150】本発明による1つの単語に対するHMMのパラメータ推定の実際の計算手順は次のようにになる。ただし、全単語に対する訓練パターン集合から、既に、ファジィクラスタリングは行われており、各クラスタ C_1, \dots, C_n のセントロイド c_1, \dots, c_n は算出済みであるとする。

【0151】(1) $L_1 = \infty$
(2) $i = 1 \sim I, j = 1 \sim I+1, m = 1 \sim M$ について
 $\lambda_i = [\{\pi_i\}_{i=1, \dots, I}, \{a_{ij}\}_{j=1, \dots, I+1}, \{b_i(C_n)\}_{n=1, \dots, n}]$
に適当な初期値を与える。

【0152】(3) $r = 1 \sim R, t = 2 \sim T^{(r)}, m = 1, \dots, M$ について
 $u(y^{(r)}(t), C_n)$ を求める。

【0153】(4) $r = 1 \sim R, t = 2 \sim T^{(r)}, i = 1 \sim I+1$ について
(数41)に従って $\omega_i(y^{(r)}(t))$ を求める。

【0154】(5) $r = 1 \sim R, t = 2 \sim T^{(r)}, i = 1 \sim I, j = 1 \sim I+1$ について
 $\alpha^{(r)}_i(t), \beta^{(r)}_i(t)$ を(数50), $\xi^{(r)}_{i,j}(t), \gamma^{(r)}_i(t)$ を(数52)に従って計算する。

【0155】(6) $i = 1 \sim I$ について
(数21)に従って π_i の再推定値 π_i^* を求める。

【0156】(7) $i = 1 \sim I, j = 1 \sim I+1$ について

(数23)に従って a_{ij} の再推定値 a_{ij}^* を求める。

【0157】(8) $i = 1 \sim I, m = 1 \sim M$ について
(数49)に従って $b_i(C_n)$ の再推定値 $b_i(C_n)^*$ を求める。

【0158】(9) $i = 1 \sim I, j = 1 \sim I+1, m = 1 \sim M$ について
 $\pi_i = \pi_i^*, a_{ij} = a_{ij}^*, b_i(C_n) = b_i(C_n)^*$ なる代入を行うことによって、再推定されたパラメータ集合 $\lambda = \{\lambda_i\}$ を得る。

【0159】(10) step (9)で得たパラメータ集合

【数50】

$$\alpha^{(r)}_i(t) = \sum_j \alpha^{(r)}_i(t-1) a_{ij} \omega_j(y^{(r)}(t)),$$

$$\beta^{(r)}_i(t) = \sum_j \omega_j(y^{(r)}(t)) a_{ij} \beta^{(r)}_j(t).$$

【0149】

【数51】

λ に対して

【0160】

【数52】

$$L_2 = \sum_{r=1}^R P(O^{(r)} | \lambda) = \sum_{r=1}^R \alpha^{(r)}_{1,1} (T^{(r)} + 1)$$

【0161】を計算する。

(9) $|L_1 - L_2| / L_1 > \epsilon$ ならば、 $L_2 = L_1$ とおいてステップ(4)へ、そうでなければ終了。

【0162】(図1)は、本発明のHMM作成装置の一実施例である。以下図面に従って説明する。

【0163】101は特徴抽出部であって、周知の方法によって、単語 w ($= 1, \dots, W$)に対応するモデル作成のために準備された訓練単語 $r = 1 \sim R$ の音声信号を特徴ベクトルの系列 $Y^{(r)} = (y^{(r)}(1), y^{(r)}(2), \dots, y^{(r)}(T^{(r)}))$ に変換するものである。

【0164】102は単語パターン記憶部であって、モデル λ^* を作成するための学習用単語を前記特徴ベクトル系列の形で R 個記憶するものである。

【0165】103はファジィクラスタリング部であって、訓練ベクトル集合から(数32) (数33)に従ってファジィクラスタリングするものである。

【0166】104はセントロイド記憶部であって、前記ファジィクラスタリングの結果得られた各クラスタのセントロイドが記憶される。このセントロイドは入力ベクトルの各クラスタに対する帰属度を算出するのに用いられる。

【0167】105はバッファメモリであって、単語パターン記憶部102に記憶されている w に対する単語パターンを R 個取り出して一時的に記憶するものである。

【0168】106はベクトル帰属度算出・記憶部であって、セントロイド記憶部104に記憶されているセントロイドからバッファメモリ105の出力ベクトルの各クラスタに対する帰属度を(数36)に従って算出するものである。

【0169】107はパラメータ推定部であって、前記

モデル λ^w を作成するステップ (1) ~ (10) を実行し、単語 w に対応するモデル λ^w を推定するものである。

【0170】 108 は第 1 のパラメータ記憶部であって、前記ステップ (9) で得られたパラメータの再推定値を一次的に記憶するものである。パラメータ推定部 107 はこのパラメータ記憶部 108 の値を用いて再推定を行うものである。

【0171】 109 は単語 $w = 1 \sim W$ に対応するパラメータを記憶する第 2 のパラメータ記憶部であって、前記それぞれの単語 $w = 1, \dots, W$ に対応するパラメータが、パラメータ記憶部 1, \dots, パラメータ記憶部 W にそれぞれ記憶される。即ち、それぞれの単語の各状態に対応する遷移確率は、第 1 のパラメータ記憶部 108 から読み出され、 w, i, j で参照可能な形で記憶される。

【0172】 以上のようにして、FVQHMM が作成される。次に、以上のようなモデルを用いて実際の入力音声を認識する方法及び装置について説明する。

【0173】 (図 4) はその認識装置のブロック図である。以下、本図に従って説明する。401 は特徴抽出部であって、(図 1) 101 と全く同様の構成・機能を有するものである。

【0174】 402 はセントロイド記憶部であって、(図 1) の HMM 作成装置のセントロイド記憶部に記憶されている各クラスタのセントロイドが記憶されている。

$$\log \omega_i^w(y(t)) = \sum_{n=1}^N u(y(t), C_n) \log b_{i,n}(C_n)$$

あるいは、

$$\omega_i^w(y(t)) = \exp \left(\sum_{n=1}^N u(y(t), C_n) \log b_{i,n}(C_n) \right)$$

【0179】 で与え、(数 5) における $b_{i,n}(o(t))$ を $y_i^w(t)$ 、 a_{ij} を a_{ij} として、(数 1) (数 2) (数 3) 等の何れかを用いて実行される。(数 1) を計算する場合は、(数 2.6) における $Y^{(t)}$ に対する

【0180】

【数 5.4】

$$\alpha_{i+1}^{(t)}(T^{(t)} + 1)$$

【0181】 の計算と全く同様に、入力パターン Y に対する $\alpha_{i+1}^{(t)}(T + 1)$ を計算することになる。T は Y のフレーム数である。

【0182】 (数 2) (数 3) を用いる場合は、周知の Viterbi 法によって尤度を求めることが出来る。漸化式計算が足し算で済み、演算の途中でアンダーフロー等の生じる心配の無い (数 3) を用いる場合が普通である。

$$\phi_{i+1}^{(t)}(t) = \max_{i=1, \dots, I} [\phi_{i+1}^{(t-1)}(t-1) + \log a_{i+1}^{(t)} + \log \omega_i^w(y(t))]$$

$$(3) \phi_{i+1}^{(t)}(T+1) = \max_{i=1, \dots, I} [\phi_{i+1}^{(t)}(T) + \log a_{i+1, T+1}^{(t)}]$$

【0187】 ステップ (3) における $\phi_{i+1}^{(t)}(T+1)$ が Y に対するモデル w (単語 w) の尤度である。

る。

【0175】 403 はベクトル帰属度算出部であって、特徴抽出部 401 の出力の特徴ベクトル $y(t)$ とセントロイド記憶部 403 に記憶されている前記それぞれのクラスタの代表ベクトル c_m ($m = 1, \dots, M$) から、 $y(t)$ をファジィベクトル量子化するものである。即ち、(数 3.3) から $y(t)$ のクラスタ C_m に対する帰属度 $u(y(t), C_m)$ ($m = 1, \dots, M$) を算出する。即ち、 $y(t)$ は帰属度ベクトル $(u(y(t), C_1), \dots, u(y(t), C_M))^\top$ に変換される。

【0176】 404 はパラメータ記憶部であって、(図 1) 109 と全く同様の構成・機能を有するものであって、パラメータ記憶部 w には、単語 w ($= 1, \dots, W$) に対応するモデルのパラメータ、 $\pi_{i,j}$, $a_{i,j}$, $b_{i,j}(C_m)$ が記憶されている。

【0177】 405 は尤度計算部であって、ベクトル帰属度算出部 403 の出力に得られる帰属度ベクトル列に対する各モデルの尤度をパラメータ記憶部 404 の内容を用いて計算するものである。即ち、尤度計算部 w ではパラメータ記憶部 w の内容が用いられる。尤度の計算は、モデル λ^w の状態 i における $y(t)$ の発生度合 $\omega_i^w(y(t))$ を(数 4.1) に準じて

【0178】

【数 5.3】

で、ここでも、(数 3) を用いる場合について説明する。

【0183】 (1) 初期値設定

単語 w の状態 i の初期確率を $\pi_{i,0}$ とし、 $i = 1, \dots, I$ について(数 5.5) を実行

【0184】

【数 5.5】

$$\phi_{i,0}^{(1)} = \log \pi_{i,0} + \log \omega_i^w(y(1))$$

【0185】 (2) 漸化式の計算

$t = 2, \dots, T$, $j = 1, \dots, I$ について(数 5.6) を実行

【0186】

【数 5.6】

【0188】 406 は比較判定部であって、尤度計算部 405 に含まれる尤度計算部 $1, \dots, W$ の何れの出力が

最大であるかを比較判定し、それに対応する単語を認識結果として出力する。即ち、

【0189】

【数57】

$$w^* = \operatorname{argmax}_{w=1, \dots, N} [\phi^w]_{i+1}(T+1)$$

【0190】を見出すものであって、これは(数4)に相当する計算である。

【0191】

【発明の効果】以上のように、本発明は、学習に用いたパターン集合を形成するベクトルの集合をファジィクラスタリングし、クラスタ C_i の前記HMMの状態 i におけるクラスタ C_i の発生確率 $b_i(C_i)$ と、入力ベクトル $y(t)$ の各クラスタへの帰属度とから $y(t)$ の状態 i における発生確率を算出するものであって、該発生確率の対数値を前記 $b_i(C_i)$ の対数値の前記帰属度の荷重和または荷重平均で表すところに特徴がある。このようにすることによって、離散型HMMにおける欠点である訓練データの不足やその偏りによる推定誤差を解消し、離散型HMMのもつ計算量が少ないというFVQHMMが本来有している利点をもつモデルを、数学的に矛盾の無い形で実現することが可能となった。

【0192】なお、本実施例においては、単語を認識す

るとして述べたが、単語を音韻や音節等に置き換えるても勿論よく、また、本発明は音声以外のパターンにも適用出来るものである。

【図面の簡単な説明】

【図1】本発明によるHMMのパラメータ推定を行う装置の一実施例を示すブロック図

【図2】HMMを用いた音声認識装置の従来例を説明するブロック図

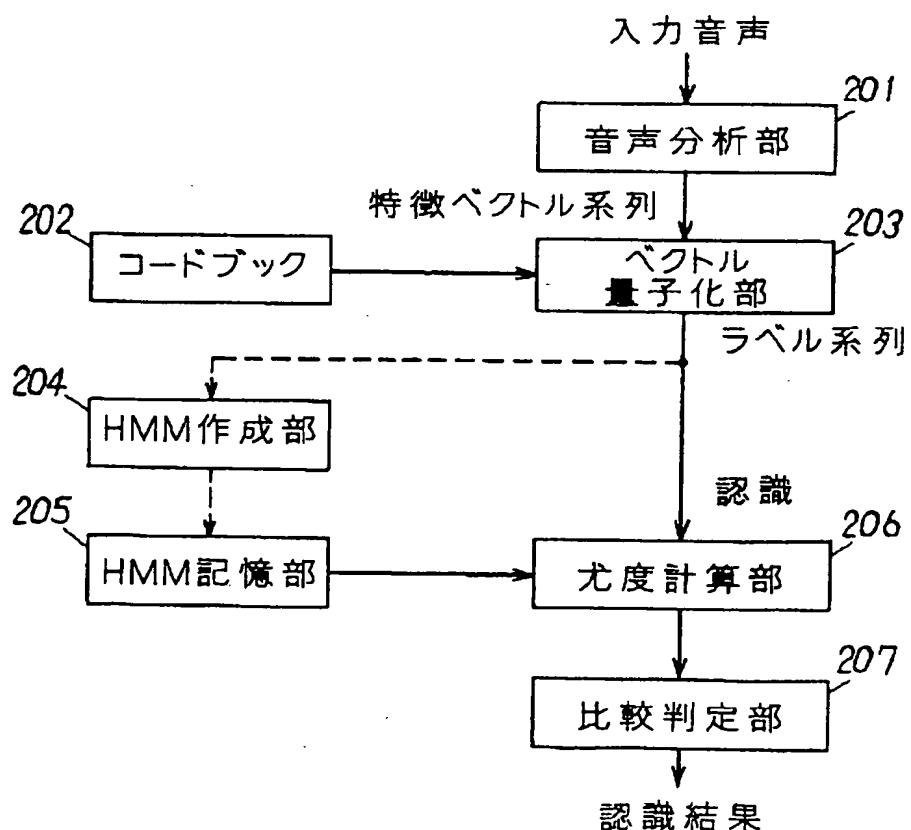
【図3】離散確率分布型HMMの構成を示すHMMの構成図

【図4】本発明の原理により構成されたHMMを用いた音声認識装置の一実施例を示すブロック図

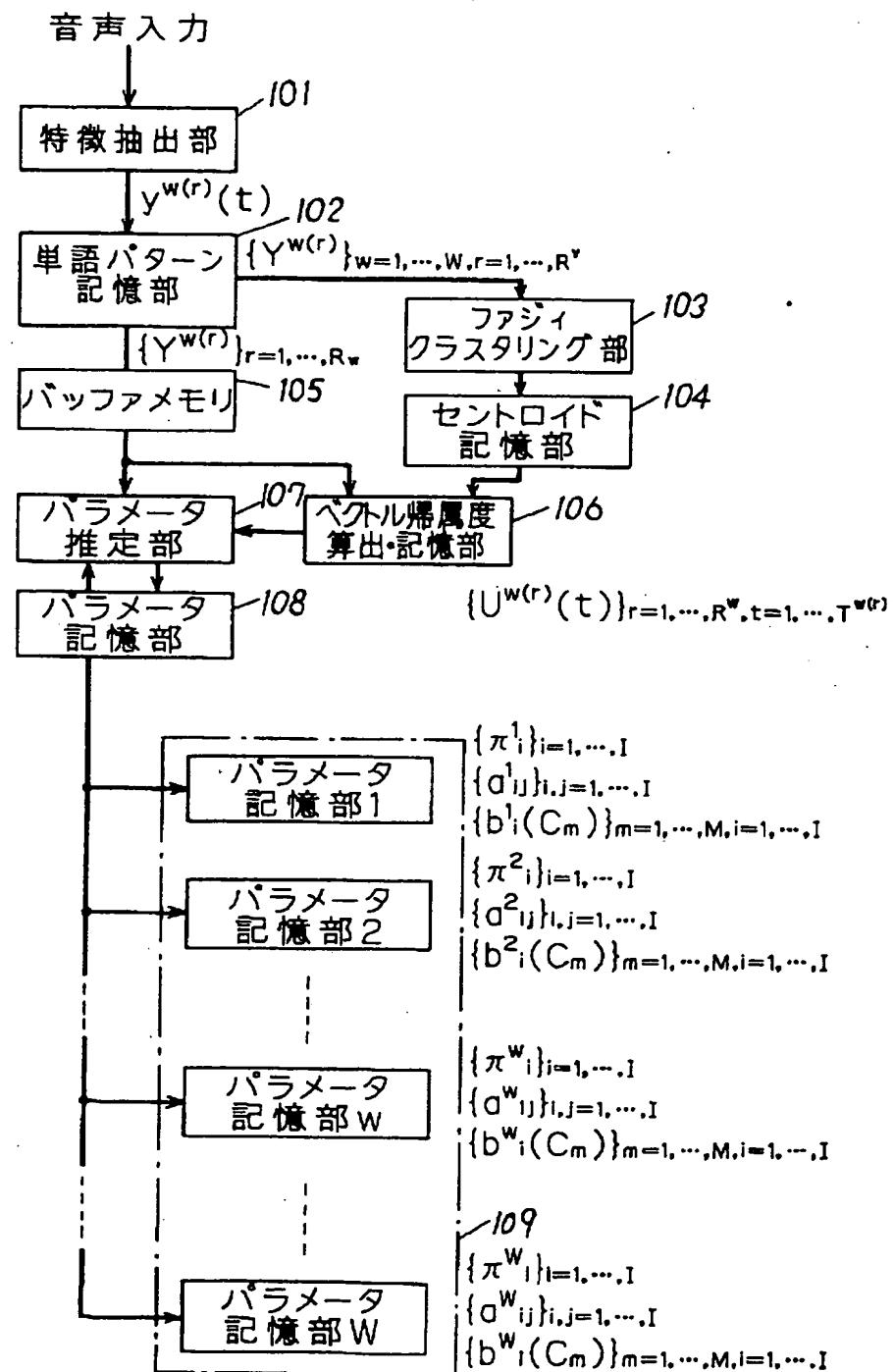
【符号の説明】

- 101 特徴抽出部
- 102 単語パターン記憶部
- 103 ファジィクラスタリング部
- 104 セントロイド記憶部
- 105 バッファメモリ
- 106 ベクトル帰属度・算出記憶部
- 107 パラメータ推定部
- 108 パラメータ記憶部
- 109 パラメータ記憶部

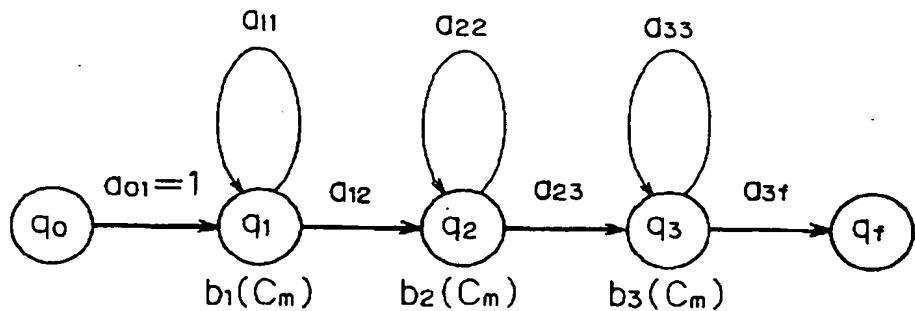
【図2】



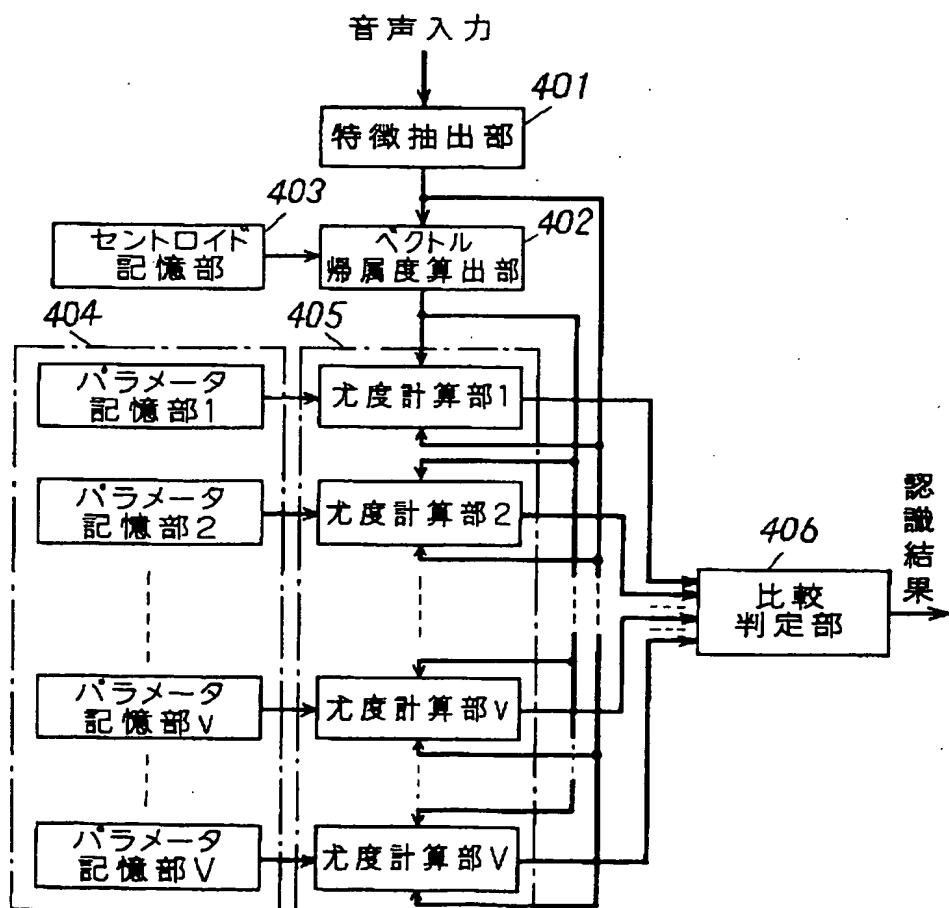
【図1】



【図3】



【図4】



【公報種別】特許法第17条の2の規定による補正の掲載

【部門区分】第6部門第2区分

【発行日】平成10年(1998)7月31日

【公開番号】特開平5-333898

【公開日】平成5年(1993)12月17日

【年通号数】公開特許公報5-3339

【出願番号】特願平4-142399

【国際特許分類第6版】

G10L 3/00 535

521

【F I】

G10L 3/00 535

521 C

【手続補正書】

【提出日】平成8年11月13日

【補正内容】

【手続補正1】

【0021】

【補正対象書類名】明細書

【数7】

【補正対象項目名】0021

【補正方法】変更

$$L_2^*(w) = \max_x [P(Y, X | \lambda^w)]$$

【手続補正2】

【補正内容】

【補正対象書類名】明細書

【0039】

【補正対象項目名】0039

【数9】

【補正方法】変更

$$O(t) = \underset{\substack{m \in \{1, 2, \dots, M\} \\ m=1, 2, \dots, M}}{\operatorname{argmin}} [d(y(t), c_m)]$$

【手続補正3】

【手続補正4】

【補正対象書類名】明細書

【補正対象書類名】明細書

【補正対象項目名】0045

【補正対象項目名】0048

【補正方法】変更

【補正方法】変更

【補正内容】

【補正内容】

【0045】

【0048】

【数11】

【数12】

$$c_m = \frac{1}{|S_m|} \sum_{n: v_n \in S_m} v_n.$$

$$D(i) = \sum_{m=1}^M \sum_{n: v_n \in S_m} d(c_m, v_n)$$

【手続補正5】

【補正方法】変更

【補正対象書類名】明細書

【補正内容】

【補正対象項目名】0055

【0055】

【数14】

$$Q(\lambda, \lambda')$$

$$= \sum_{x^{(1)}} \cdots \sum_{x^{(R)}} \left\{ \prod_{k=1}^R P(O^{(k)}, X^{(k)} | \lambda) \right\}$$

$$\log \left\{ \prod_{k=1}^R P(O^{(k)}, X^{(k)} | \lambda) \right\}$$

【手続補正6】

【補正対象書類名】明細書

【補正対象項目名】0081

【補正方法】変更

【補正内容】

【0081】

【数26】

$$\alpha^{(r)}_i(t) = P(o^{(r)}(1), \dots, o^{(r)}(t), X^{(r)}(t) = i | \lambda)$$

$$\beta^{(r)}_i(t) = P(o^{(r)}(t+1), \dots, o^{(r)}(T) | X^{(r)}(t) = i, \lambda)$$

【手続補正7】

【補正対象書類名】明細書

【補正対象項目名】0106

【補正方法】変更

【補正内容】

【0106】

【数34】

$$U_{mn} = \frac{1}{\sum_{k=1}^M \left(\frac{d_{mn}}{d_{kn}} \right)^{1/(F-1)}} \quad 1 \leq m \leq M, \quad 1 \leq n \leq N.$$

【手続補正8】

【補正対象書類名】明細書

【補正対象項目名】0108

【補正方法】変更

【補正内容】

【0108】

【数35】

$$(d_{mn}/d_{kn})^{1/(F-1)} \rightarrow \begin{cases} 1 & k = m \text{ のとき} \\ 0 & k \neq m, d_{mn} < d_{kn} \text{ のとき} \\ \infty & k \neq m, d_{mn} > d_{kn} \text{ のとき} \end{cases}$$

【手続補正9】

【補正対象書類名】明細書

【補正対象項目名】0110

【補正方法】変更

【補正内容】

【0110】

【数36】

$$U_{mn} \rightarrow \begin{cases} 1 & d_{mn} \text{が最小のとき} \\ 0 & d_{mn} \text{が最小でないとき} \end{cases}$$

【手続補正10】

【補正対象書類名】明細書

【補正対象項目名】0118

【補正方法】変更

【補正内容】

【0118】

【数37】

$$\omega_i(y^{(r)}(t)) = \sum_{m=1}^M u(y^{(r)}(t), C_m) b_i(C_m)$$

$$u(y^{(r)}(t), C_m) = \frac{1}{\sum_{k=1}^M \left[\frac{d(y^{(r)}(t), C_m)}{d(y^{(r)}(t), C_k)} \right]^{1/(F-1)}}$$

【手続補正11】

【補正対象書類名】明細書

【補正対象項目名】0127

【補正方法】変更

【補正内容】

【0127】

【数41】

$$\omega_i(y^{(r)}(t)) = \prod_{m=1}^M \left[(b_i(C_m))^{u(y^{(r)}(t), C_m)} \right]$$

即ち

$$\log \omega_i(y^{(r)}(t)) = \sum_{m=1}^M u(y^{(r)}(t), C_m) \log b_i(C_m).$$

【手続補正12】

【補正対象書類名】明細書

【補正対象項目名】0139

【補正方法】変更

【補正内容】

【0139】

【数46】

$$0 = \frac{\partial}{\partial b_i(C_m)} \left\{ Q_3 - \left[\sum_{k=1}^M b_i(C_k) - 1 \right] \right\}$$

$$= \sum_{r=1}^R h^{(r)} \sum_t r^{(r)}(t) u(0^{(r)}(t), C_m) \frac{1}{b_i(C_m)} - \theta$$

【手続補正13】

【補正対象書類名】明細書

【補正対象項目名】0178

【補正方法】変更

【補正内容】

【0178】

【数53】

$$\log \omega^{w_i}(y(t)) = \sum_{m=1}^M u(y(t), C_m) \log b^{w_i}(C_m)$$

あるいは、

$$\omega^{w_i}(y(t)) = \exp \left[\sum_{m=1}^M u(y(t), C_m) \log b^{w_i}(C_m) \right]$$

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- BLACK BORDERS**
- IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- FADED TEXT OR DRAWING**
- BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- SKEWED/SLANTED IMAGES**
- COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- GRAY SCALE DOCUMENTS**
- LINES OR MARKS ON ORIGINAL DOCUMENT**
- REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.